

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Schmidt, Eduard Daniel Leendert
De Vries, Sape Cornelis
Hecht, Valerie France Gabrielle
- (ii) TITLE OF INVENTION: Apomixis Conferred By Expression of SERK
Interacting Proteins
- (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Syngenta Patent Dept.
 - (B) STREET: 3054 Cornwallis Road
 - (C) CITY: RTP
 - (D) STATE: NC
 - (E) COUNTRY: USA
 - (F) ZIP: 27709
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: PCT/EP99/07972
 - (B) FILING DATE: 20-OCT-1999
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Meigs, J. Timothy
 - (B) REGISTRATION NUMBER: 38,241
 - (C) REFERENCE/DOCKET NUMBER: S-30683A
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 919-541-8587

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 551 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

(B) CLONE: 3A35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACGTGTCGGT GGAGCGGGT CGGTCAGTC GGGTCAGATA CCAAGGTGCC AAGTGAAGG	60
TTGTGGGATG GATCTAACCA ATGCAAAAGG TTATTACTCG AGACACCGAG TTTGTGGAGT	120
GCACTCTAAA ACACCTAAAG TCACTGTGGC TGGTATCGAA CAGAGGTTTT GTCAACAGTG	180
CAGCAGGTTT CATCAGCTTC CGGAATTTGA CCTAGAGAAA AGGAGTTGCC GCAGGAGACT	240
CGCTGGTCAT AATGAGCGAC GAAGGAAGCC ACAGCCTGCG TCTCTCTCTG TGTTAGCTTC	300
TCGTTACGGG AGGATCGCAC CTTCGCTTTA CGAAAATGGT GATGCTGGAA TGAATGGAAG	360
CTTTCTTGGG AACCAAGAGA TAGGATGGCC AAGTTCAAGA ACATTGGATA CAAGAGTGAT	420
GAGGCGGCCA GTGTCATCAC CGTCATGGCA GATCAATCCA ATGAATGTAT TTAGTCAAGG	480
TTCAGTTGGT GGAGGAAGGA CAAGCTTCTC ATCTCCAGAG ATTATGGACA CTAAACTAGA	540
GAGCTACAAG G	551

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 375 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

(B) CLONE: 3A35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Glu Met Gly Ser Asn Ser Gly Pro Gly His Gly Pro Gly Gln Ala
 1 5 10 15

Glu Ser Gly Gly Ser Ser Thr Glu Ser Ser Ser Phe Ser Gly Gly Leu
 20 25 30

Met Phe Gly Gln Lys Ile Tyr Phe Glu Asp Gly Gly Gly Gly Ser Gly
 35 40 45

Ser Ser Ser Ser Gly Gly Arg Ser Asn Arg Arg Val Arg Gly Gly Gly
 50 55 60

Ser Gly Gln Ser Gly Gln Ile Pro Arg Cys Gln Val Glu Gly Cys Gly
 65 70 75 80

Met Asp Leu Thr Asn Ala Lys Gly Tyr Tyr Ser Arg His Arg Val Cys
 85 90 95

Gly Val His Ser Lys Thr Pro Lys Val Thr Val Ala Gly Ile Glu Gln
 100 105 110

Arg Phe Cys Gln Gln Cys Ser Arg Phe His Gln Leu Pro Glu Phe Asp
 115 120 125

Leu Glu Lys Arg Ser Cys Arg Arg Arg Leu Ala Gly His Asn Glu Arg
 130 135 140

Arg Arg Lys Pro Gln Pro Ala Ser Leu Ser Val Leu Ala Ser Arg Tyr
 145 150 155 160

Gly Arg Ile Ala Pro Ser Leu Tyr Glu Asn Gly Asp Ala Gly Met Asn
 165 170 175

Gly Ser Phe Leu Gly Asn Gln Glu Ile Gly Trp Pro Ser Ser Arg Thr
 180 185 190

Leu Asp Thr Arg Val Met Arg Arg Pro Val Ser Ser Pro Ser Trp Gln
 195 200 205

Ile Asn Pro Met Asn Val Phe Ser Gln Gly Ser Val Gly Gly Gly Arg
 210 215 220

Thr Ser Phe Ser Ser Pro Glu Ile Met Asp Thr Lys Leu Glu Ser Tyr

225		230		235		240									
Lys	Gly	Ile	Gly	Asp	Ser	Asn	Cys	Ala	Leu	Ser	Leu	Leu	Ser	Asn	Pro
				245					250					255	
His	Gln	Pro	His	Asp	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Asn
				260					265					270	
Asn	Asn	Asn	Thr	Trp	Arg	Ala	Ser	Ser	Gly	Phe	Gly	Pro	Met	Thr	Val
			275					280					285		
Thr	Met	Ala	Gln	Pro	Pro	Pro	Ala	Pro	Ser	Gln	His	Gln	Tyr	Leu	Asn
			290					295				300			
Pro	Pro	Trp	Val	Phe	Lys	Asp	Asn	Asp	Asn	Asp	Met	Ser	Pro	Val	Leu
			305				310				315				320
Asn	Leu	Gly	Arg	Tyr	Thr	Glu	Pro	Asp	Asn	Cys	Gln	Ile	Ser	Ser	Gly
				325					330						335
Thr	Ala	Met	Gly	Glu	Phe	Glu	Leu	Ser	Asp	His	His	His	Gln	Ser	Arg
				340					345					350	
Arg	Gln	Tyr	Met	Glu	Asp	Glu	Asn	Thr	Arg	Ala	Tyr	Asp	Ser	Ser	Ser
				355				360					365		
His	His	Thr	Asn	Trp	Ser	Leu									
				370				375							

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 859 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Arabidopsis thaliana
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: 3B39

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TCAACATGTC TTCCTAACCA GAAATCCACC ATCATCTTCC CACGAATACA ACTTAAAGCT	60
TTACCAGAAA ATGGAGGGTC AGAGAACACA ACGCCGGGGT TACTTGAAAG ACAAGGCTAC	120
AGTCTCCAAC CTTGTTGAAG AAGAAATGGA GAATGGCATG GATGGAGAAG AGGAGGATGG	180
AGGAGACGAA GACAAAAGGA AGAAGGTGAT GGAAAGAGTT AGAGGTCCTA GCACTGACCG	240
TGTTCCATCG CGACTGTGCC AGGTCGATAG GTGCACTGTT AATTTGACTG AGGCCAAGCA	300
GTATTACCGC AGACACAGAG TATGTGAAGT ACATGCAAAG GCATCTGCTG CGACTGTTGC	360
AGGGGTCAGG CAACGCTTTT GTCAACAATG CAGCAGGTTT CATGAGCTAC CAGAGTTTGA	420
TGAAGCTAAA AGAAGCTGCA GGAGGCGCTT AGCTGGACAC AATGAGAGGA GGAGGAAGAT	480
CTCTGGTGAC AGTTTTGGAG AAGGGTCAGG CCGGAGAGGG TTTAGCGGTC AACTGATCCA	540
GACTCAAGAA AGAAACAGGG TAGACAGGAA ACTTCCTATG ACCAACTCAT CATTTAAGGG	600
ACCACAGATC AGATAAACC TCCCGCTCTC TCTCTTCTGT CATCTACATA TGCTCTATCT	660
ACACTCTTAT TAGACAAATA ATGGCATCTA ACAATGTCAA GAAAAGTTGG TCATGGTATT	720
AAATCCTAGA GGGAAATATA AGTATAAACC TTTAGTCCCC TTTATGCTGT CCTGTAATGA	780
ATATCTATCC GGAAATGTAT TCGCATAGTC TTGCGTCTAA TAATGTTTAT TAAAAAAAAA	840
AAAAAAAAA AAAAAAAAAA	859

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 181 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*

(B) CLONE: 3B39

[illegible]

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 479 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

(B) CLONE: 4B19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AGAAGCAGAA AGGTAAAGCT ACAAGTAGTA GTGGAGTTTG TCAGGTCGAG AGTTGTACCG	60
CGGATATGAG CAAAGCCAAA CAGTACCACA AACGACACAA AGTCTGCCAG TTTCATGCCA	120
AAGCTCCTCA TGTTCCGATC TCTGGTCTTC ACCAACGTTT CTGCCAACAA TGCAGCAGGT	180
TTCACGCGCT CAGTGAGTTT GATGAAGCCA AGCGGAGTTG CAGGAGACGC TTAGCTGGAC	240
ACAACGAGAG AAGGCGGAAA AGCACAACCTG ACTAAAGACG GTGAAACGTG TGAGATCCCG	300
GTTTGAAGGT TAATGAAACA GGCTTTGCTT ACTCTCTTCT GTCAGTCTCT TTTAGCTCCT	360
TGTAATCCTC TGIGTCTCTG TCTGTTTCTC CATATTACCT GTAATCAAAG CTATCTGCTA	420
AACCTACGAC ATGGTTAAAT AAATGCATTG AGACTTAAAA AAAAAAAAAA AAAAAAAAAA	479

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 131 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

(B) CLONE: 4B19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Ser Met Arg Arg Ser Lys Ala Glu Gly Lys Arg Ser Leu Arg Glu
1 5 10 15

Leu Ser Glu Glu Glu Glu Glu Glu Glu Glu Thr Glu Asp Glu Asp Thr
20 25 30

Phe Glu Glu Glu Glu Ala Leu Glu Lys Lys Gln Lys Gly Lys Ala Thr
35 40 45

Ser Ser Ser Gly Val Cys Gln Val Glu Ser Cys Thr Ala Asp Met Ser
50 55 60

Lys Ala Lys Gln Tyr His Lys Arg His Lys Val Cys Gln Phe His Ala
65 70 75 80

Lys Ala Pro His Val Arg Ile Ser Gly Leu His Gln Arg Phe Cys Gln
85 90 95

Gln Cys Ser Arg Phe His Ala Leu Ser Glu Phe Asp Glu Ala Lys Arg
100 105 110

Ser Cys Arg Arg Arg Leu Ala Gly His Asn Glu Arg Arg Arg Lys Ser
115 120 125

Thr Thr Asp
130

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2682 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

(B) CLONE: 3A52

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GCCATTCAAG GAGACACTAA TGGTGCTCTT ACTTTGAATC TTAATGGTGA AAGTGATGGC	60
CTTTTTCCTG CCAAGAAGAC CAAATCCGGA GCCGTTTGTC AGGTCGAAAA CTGTGAAGCT	120
GATCTTAGTA AAGTTAAGGA TTATCATAGA CGCCATAAGG TCTGTGAGAT GCATTCCAAG	180
GCTACTAGTG CCACTGTGGG AGGTATCTTG CAGCGCTTTT GTCAGCAATG TAGTAGGTTT	240
CATCTTCTGC CAGGTTTCGA TGACGGAAAG AGAAGTTGTC GTAGACGTTT GGCTGGCCAT	300
AATAAACGTC CGAGGAAAC AAATCCCGAA CCTGGCGCTA ACGGGAATCC TAGTGATGAT	360
CACTCAAGCA ACTATCTCTT GATTACTCTC TTGAAGATAC TCTCCAATAT GCATAACCAT	420
ACCGGTGATC AAGATTTGAT GTCTCATCTT CTGAAGAGCC TCGTAAGCCA TGCTGGCGAA	480
CAGTTAGGGA AAAACTTAGT TGAACCTCTT CTACAAGGAG AGATCTCAAG GTTCCTTAAA	540
ATATTGGAAA ACTCGGCTTT GCTTGGGATT GAGCAAGCTC CTCAAGAGGA GTTAAAGCAA	600
TTTTGGGCTC GGCAAGATGG GACAGCTACC GAGAACAGAT CAGAAAAACA AGTCAAAATG	660
AATGATTTTG ATTTGAATGA TATCTATATA GACTCAGATG ACACAGACGT CGAAAGATCT	720
CCTCCTCCAA CGAATCCAGC GACCAGTTCT CTTGATTATC CTTCATGGAT ACATCAGTCT	780
AGTCCGCCTC AGACAAGTAG GAATTCAGAT TCAGCATCTG ACCAGTCACC CTCAAGTTCT	840
AGTGAAGATG CTCAGATGCG CACAGGCCGG ATTGTGTTCA AACTATTTGG GAAAGAGCCA	900
AATGAATTTT CTATTGTCTT ACGAGGACAG ATTCTTGACT GGTATATCGA TAGTCCAAC	960
GACATGGAGA GCTACATAAG ACCTGGCTGT ATCGTATTGA CCATCTATCT TCGTCAAGCT	1020
GAAACTGCTT GGAAGAACT TTCAGACGAT CTGGGTTTGA GCTTAGGGAA GCTTCTAGAT	1080
CTCTCCGATG ATCCCTTGTT GACAACTGGA TGGATTTATG TAGGGTGCAG AACCAACTTG	1140
CATTGTGATA TAACGGTCAG GTTGTGTTG ACACTTCATT GTCTCTAAAA AGTCGTGATT	1200
ATAGTCACAT CATTAGCGTT AAACCGCTTG CTATAGCTGC AACGGAGAAG GCTCAATTTA	1260

CAGTTAAAGG	CATGAATCTC	CGTCGGCGTG	GCACAAGGTT	ACTTTGTTCT	GTTGAAGGAA	1320
AATACTTGAT	TCAGGAAACA	ACACACGATT	CGACGACCAG	GGAGGATGAC	GATTTCAAGG	1380
ACAACAGTGA	GATTGTTGAG	TGTGTAAACT	TCTCTTGTA	TATGCCTATA	TTGAGTGGTC	1440
GAGGATTCAT	GGAGATTGAA	GACCAAGGAC	TCAGTAGCAG	CTTCTTCCCT	TTCTTAGTGG	1500
TTGAAGATGA	CGATGTTTGT	TCTGAAATCC	GTATACTTGA	AACCACATTA	GAGTTCACCTG	1560
GAAC TGATT C	TGCTAAGCAA	GCTATGGATT	TCATACATGA	AATCGGTTGG	CTTCTTCACA	1620
GAAGTAAACT	TGGGAATCA	GACCCAAATC	CAGGCGTTTT	CCCATTAAATA	CGCTTCCAGT	1680
GGCTAATCGA	GTTCTCAATG	GATCGAGAGT	GGTGGCGTGT	GATCAGAAAG	CTATTAAACA	1740
TGTTCTTTGA	TGGAGCTGTT	GGTGAATTTT	CTTCCTCCTC	TAATGCCACA	CTGT CAGAAC	1800
TGTGCCTTCT	TCACAGAGCC	GTGAGGAAAA	ACTCTAAGCC	TATGGTTGAA	ATGCTCTTGA	1860
GATATATTCC	CAAGCAACAG	AGAAACAGCT	TGTTTAGACC	CGATGCTGCT	GGTCCAGCCG	1920
GCTTAACACC	TCTTCATATT	GCAGCTGGTA	AAGACGGTTC	AGAAGATGTG	TTGGATGCGC	1980
TAACAGAAGA	TCCTGCAATG	GTGGGGATTG	AAGCGTGGAA	GACATGTGGA	GACAGCACAG	2040
GCTTCACACC	AGAAGACTAC	GCACGCTTAC	GCGGTCACCT	CTCATAATC	CACCTGATT C	2100
AACGCAAGAT	CAATAAAAAG	TCAACAAC TG	AAGATCATGT	TGTGGTCAAC	ATCCCAGTTT	2160
CTTTCTCAGA	CAGAGAGCAG	AAAGAACCAA	AATCAGGTCC	GATGGCTTCA	GCCTTGGAGA	2220
TCACACAGAT	TCCATGCAAG	CTCTGTGACC	ATAAACTGGT	GTATGGGACA	ACACGCAGGT	2280
CTGTAGCGTA	CAGACCAGCT	ATGTTGTCAA	TGGTGGCGAT	TGCTGCGGTT	TGCGTCTGTG	2340
TGGCACTTCT	GTTTAAGAGT	TGCCCGGAAG	TGCTCTATGT	GTTTCAACCG	TT CAGGTGGG	2400
AGTTATTGGA	CTATGGAACA	AGCTGAGTGT	AAGTCTACTT	TGAAAGATCT	TCTAAGATAT	2460
ATATATGAAT	GTTACTTATA	TAAAACCATA	GAGGTGTGAT	TTCTATATGT	AACTATATGA	2520
GTATAAGATA	TAGAGACATG	TTGGAGAAGA	AGATTGTTGT	TATTATTGTT	GTTGTTGTTG	2580
TTGTGTAAAA	GCCTCTCCTA	TCTCTCTCGA	ACCTAAGGAT	TCTCTCTCTG	ATTAGTATAT	2640
TTTTTGTTTG	ACAAAAAAA	AAAAAAA	AAAAAAA	AA		2682

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 848 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

- (B) CLONE: 3A52

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```

Met Glu Ala Arg Ile Asp Glu Gly Gly Glu Ala Gln Gln Phe Tyr Gly
 1             5             10             15

Ser Val Gly Asn Ser Ser Asn Ser Ser Ser Ser Cys Ser Asp Glu Gly
          20             25             30

Asn Asp Lys Lys Arg Arg Ala Val Ala Ile Gln Gly Asp Thr Asn Gly
          35             40             45

Ala Leu Thr Leu Asn Leu Asn Gly Glu Ser Asp Gly Leu Phe Pro Ala
          50             55             60

Lys Lys Thr Lys Ser Gly Ala Val Cys Gln Val Glu Asn Cys Glu Ala
          65             70             75             80

Asp Leu Ser Lys Val Lys Asp Tyr His Arg Arg His Lys Val Cys Glu
          85             90             95

Met His Ser Lys Ala Thr Ser Ala Thr Val Gly Gly Ile Leu Gln Arg
          100            105            110

Phe Cys Gln Gln Cys Ser Arg Phe His Leu Leu Pro Gly Phe Asp Asp
          115            120            125

Gly Lys Arg Ser Cys Arg Arg Arg Leu Ala Gly His Asn Lys Arg Pro
          130            135            140

```

Arg Lys Thr Asn Pro Glu Pro Gly Ala Asn Gly Asn Pro Ser Asp Asp
 145 150 155 160
 His Ser Ser Asn Tyr Leu Leu Ile Thr Leu Leu Lys Ile Leu Ser Asn
 165 170 175
 Met His Asn His Thr Gly Asp Gln Asp Leu Met Ser His Leu Leu Lys
 180 185 190
 Ser Leu Val Ser His Ala Gly Glu Gln Leu Gly Lys Asn Leu Val Glu
 195 200 205
 Leu Leu Leu Gln Gly Arg Arg Ser Gln Gly Ser Leu Asn Ile Gly Asn
 210 215 220
 Ser Ala Leu Leu Gly Ile Glu Gln Ala Pro Gln Glu Glu Leu Lys Gln
 225 230 235 240
 Phe Ser Ala Arg Gln Asp Gly Thr Ala Thr Glu Asn Arg Ser Glu Lys
 245 250 255
 Gln Val Lys Met Asn Asp Phe Asp Leu Asn Asp Ile Tyr Ile Asp Ser
 260 265 270
 Asp Asp Thr Asp Val Glu Arg Ser Pro Pro Pro Thr Asn Pro Ala Thr
 275 280 285
 Ser Ser Leu Asp Tyr Pro Ser Trp Ile His Gln Ser Ser Pro Pro Gln
 290 295 300
 Thr Ser Arg Asn Ser Asp Ser Ala Ser Asp Gln Ser Pro Ser Ser Ser
 305 310 315 320
 Ser Glu Asp Ala Gln Met Arg Thr Gly Arg Ile Val Phe Lys Leu Phe
 325 330 335
 Gly Lys Glu Pro Asn Glu Phe Pro Ile Val Leu Arg Gly Gln Ile Leu
 340 345 350
 Asp Trp Leu Ser His Ser Pro Thr Asp Met Glu Ser Tyr Ile Arg Pro
 355 360 365
 Gly Cys Ile Val Leu Thr Ile Tyr Leu Arg Gln Ala Glu Thr Ala Trp
 370 375 380
 Glu Glu Leu Ser Asp Asp Leu Gly Phe Ser Leu Gly Lys Leu Leu Asp
 385 390 395 400

Leu Ser Asp Asp Pro Leu Trp Thr Thr Gly Trp Ile Tyr Val Arg Val
 405 410 415
 Gln Asn Gln Leu Ala Phe Val Tyr Asn Gly Gln Val Val Val Asp Thr
 420 425 430
 Ser Leu Ser Leu Lys Ser Arg Asp Tyr Ser His Ile Ile Ser Val Lys
 435 440 445
 Pro Leu Ala Ile Ala Ala Thr Glu Lys Ala Gln Phe Thr Val Lys Gly
 450 455 460
 Met Asn Leu Arg Arg Arg Gly Thr Arg Leu Leu Cys Ser Val Glu Gly
 465 470 475 480
 Lys Tyr Leu Ile Gln Glu Thr Thr His Asp Ser Thr Thr Arg Glu Asp
 485 490 495
 Asp Asp Phe Lys Asp Asn Ser Glu Ile Val Glu Cys Val Asn Phe Ser
 500 505 510
 Cys Asp Met Pro Ile Leu Ser Gly Arg Gly Phe Met Glu Ile Glu Asp
 515 520 525
 Gln Gly Leu Ser Ser Ser Phe Phe Pro Phe Leu Val Val Glu Asp Asp
 530 535 540
 Asp Val Cys Ser Glu Ile Arg Ile Leu Glu Thr Thr Leu Glu Phe Thr
 545 550 555 560
 Gly Thr Asp Ser Ala Lys Gln Ala Met Asp Phe Ile His Glu Ile Gly
 565 570 575
 Trp Leu Leu His Arg Ser Lys Leu Gly Glu Ser Asp Pro Asn Pro Gly
 580 585 590
 Val Phe Pro Leu Ile Arg Phe Gln Trp Leu Ile Glu Phe Ser Met Asp
 595 600 605
 Arg Glu Trp Cys Ala Val Ile Arg Lys Leu Leu Asn Met Phe Phe Asp
 610 615 620
 Gly Ala Val Gly Glu Phe Ser Ser Ser Ser Asn Ala Thr Leu Ser Glu
 625 630 635 640
 Leu Cys Leu Leu His Arg Ala Val Arg Lys Asn Ser Lys Pro Met Val
 645 650 655
 Glu Met Leu Leu Arg Tyr Ile Pro Lys Gln Gln Arg Asn Ser Leu Phe

660	665	670
Arg Pro Asp Ala Ala Gly Pro Ala Gly Leu Thr Pro Leu His Ile Ala		
675	680	685
Ala Gly Lys Asp Gly Ser Glu Asp Val Leu Asp Ala Leu Thr Glu Asp		
690	695	700
Pro Ala Met Val Gly Ile Glu Ala Trp Lys Thr Cys Arg Asp Ser Thr		
705	710	715
Gly Phe Thr Pro Glu Asp Tyr Ala Arg Leu Arg Gly His Phe Ser Tyr		
725	730	735
Ile His Leu Ile Gln Arg Lys Ile Asn Lys Lys Ser Thr Thr Glu Asp		
740	745	750
His Val Val Val Asn Ile Pro Val Ser Phe Ser Asp Arg Glu Gln Lys		
755	760	765
Glu Pro Lys Ser Gly Pro Met Ala Ser Ala Leu Glu Ile Thr Gln Ile		
770	775	780
Pro Cys Lys Leu Cys Asp His Lys Leu Val Tyr Gly Thr Thr Arg Arg		
785	790	795
Ser Val Ala Tyr Arg Pro Ala Met Leu Ser Met Val Ala Ile Ala Ala		
805	810	815
Val Cys Val Cys Val Ala Leu Leu Phe Lys Ser Cys Pro Glu Val Leu		
820	825	830
Tyr Val Phe Gln Pro Phe Arg Trp Glu Leu Leu Asp Tyr Gly Thr Ser		
835	840	845

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 576 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

(B) CLONE: 4B11

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CAGCGGAAGA GCTCACCGTT GAAGAGAGGA ATCTCCTCTC TGTTGCTTAC AAAAACGTGA	60
TCGGATCTCT ACGCGCCGCC TGGAGGATCG TGTCTTCGAT TGAGCAGAAG GAAGAGAGTA	120
GGAAGAACGA CGAGCACGTG TCGCTTGTC AAGGATTACAG ATCTAAAGTT GAGTCTGAGC	180
TTTCTTCTGT TTGCTCTGGA ATCCTTAAGC TCCTTGACTC GCATCTGATC CCATCTGCTG	240
GAGCGAGTGA GTCTAAGGTC TTTTACTTGA AGATGAAAGG TGATTATCAT CGGTACATGG	300
CTGAGTTTAA GTCTGGTGAT GAGAGGAAAA CTGCTGCTGA AGATACCATG CTCGCTTACA	360
AAGCAGCTCA GGATATCGCA GCTGCGGATA TGGCACCTAC TCATCCGATA AGGCTTGGTC	420
TGGCCCTGAA TTTCTCAGTG TTCTACTATG AGATTCTCAA TTCTTCAGAC AAAGCTTGTA	480
ACATGGCCAA ACAGGCTTTT GAGGAAGCCA TAGCTGAGCT TGACACTCTG GGAGAAGAAT	540
CCTACAAAGA CAGCACTCTC ATAATGCAGT TGCTGA	576

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 248 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

(B) CLONE: 4B11

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met	Ala	Ala	Thr	Leu	Gly	Arg	Asp	Gln	Tyr	Val	Tyr	Met	Ala	Lys	Leu	1	5	10	15
Ala	Glu	Gln	Ala	Glu	Arg	Tyr	Glu	Glu	Met	Val	Gln	Phe	Met	Glu	Gln	20	25	30	
Leu	Val	Thr	Gly	Ala	Thr	Pro	Ala	Glu	Glu	Leu	Thr	Val	Glu	Glu	Arg	35	40	45	
Asn	Leu	Leu	Ser	Val	Ala	Tyr	Lys	Asn	Val	Ile	Gly	Ser	Leu	Arg	Ala	50	55	60	
Ala	Trp	Arg	Ile	Val	Ser	Ser	Ile	Glu	Gln	Lys	Glu	Glu	Ser	Arg	Lys	65	70	75	80
Asn	Asp	Glu	His	Val	Ser	Leu	Val	Lys	Asp	Tyr	Arg	Ser	Lys	Val	Glu	85	90	95	
Ser	Glu	Leu	Ser	Ser	Val	Cys	Ser	Gly	Ile	Leu	Lys	Leu	Leu	Asp	Ser	100	105	110	
His	Leu	Ile	Pro	Ser	Ala	Gly	Ala	Ser	Glu	Ser	Lys	Val	Phe	Tyr	Leu	115	120	125	
Lys	Met	Lys	Gly	Asp	Tyr	His	Arg	Tyr	Met	Ala	Glu	Phe	Lys	Ser	Gly	130	135	140	
Asp	Glu	Arg	Lys	Thr	Ala	Ala	Glu	Asp	Thr	Met	Leu	Ala	Tyr	Lys	Ala	145	150	155	160
Ala	Gln	Asp	Ile	Ala	Ala	Ala	Asp	Met	Ala	Pro	Thr	His	Pro	Ile	Arg	165	170	175	
Leu	Gly	Leu	Ala	Leu	Asn	Phe	Ser	Val	Phe	Tyr	Tyr	Glu	Ile	Leu	Asn	180	185	190	
Ser	Ser	Asp	Lys	Ala	Cys	Asn	Met	Ala	Lys	Gln	Ala	Phe	Glu	Glu	Ala	195	200	205	
Ile	Ala	Glu	Leu	Asp	Thr	Leu	Gly	Glu	Glu	Ser	Tyr	Lys	Asp	Ser	Thr	210	215	220	
Leu	Ile	Met	Gln	Leu	Leu	Arg	Asp	Asn	Leu	Thr	Leu	Trp	Thr	Ser	Asp	225	230	235	240

Met Gln Glu Gln Met Asp Glu Ala
245

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 659 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

- (B) CLONE: 4A24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CGCCGCCACC GCGATGTACG TGATCTACCA CCCTCGTCCG CCGTCGTCTT CCGTCCCGTC	60
AATAAGAATC AGCCGCGTGA ACCTAACAAC CTCCTCTGAT TCCTCCGTCT CTCATCTCTC	120
TTCCITCTTC AACTTCACTC TAATCTCAGA GAATCCAAAC CAACACCTCT CTTTCTCTTA	180
CGATCCTTTC ACGTCCACCG TTAATTCAGC TAAATCCGGT ACGATGCTCG GTAACGGAAC	240
TGTTCCTGCT TTCTTCAGCG ATAACGGTAA CAAACTTCG TTTCACGGCG TGATCGCTAC	300
GTCTACAGCG GCGCGTGAGT TAGATCCGGA TGAAGCTAAG CATCTGAGAT CAGATCTGAC	360
GCGCGCGCGT GTAGGATATG AGATCGAGAT GAGAACTAAA GTGAAGATGA TAATGGGGAA	420
GCTGAAGAGT GAAGGAGTAG AGATCAAAGT GACATGTTGA AGGATTTGAA GGAACATATC	480
CAAAAGGTAA AACTCCAATT GTAGCTACTT CTAAAAAAC TAAGTGTAAG TCTGATCTTA	540
GTGTCAAGTC TGGAAATGGA TTTCTAAAGG AATTTGATAA TTTCACATTG AAATTCTATA	600
TATCTCTCTT TTCTCTGGA TTTGTGAAAC TTTGGATGAT CAAAGAATTC TTCATTGTC	659

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 174 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:
 (B) CLONE: 4A24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

```

Arg Ile Cys Cys Cys Cys Phe Trp Ser Ile Leu Ile Ile Leu Ile Leu
1           5           10           15

Ala Leu Met Thr Ala Ile Ala Ala Thr Ala Met Tyr Val Ile Tyr His
          20           25           30

Pro Arg Pro Pro Ser Phe Ser Val Pro Ser Ile Arg Ile Ser Arg Val
          35           40           45

Asn Leu Thr Thr Ser Ser Asp Ser Ser Val Ser His Leu Ser Ser Phe
          50           55           60

Phe Asn Phe Thr Leu Ile Ser Glu Asn Pro Asn Gln His Leu Ser Phe
          65           70           75           80

Ser Tyr Asp Pro Phe Thr Val Thr Val Asn Ser Ala Lys Ser Gly Thr
          85           90           95

Met Leu Gly Asn Gly Thr Val Pro Ala Phe Phe Ser Asp Asn Gly Asn
          100          105          110

Lys Thr Ser Phe His Gly Val Ile Ala Thr Ser Thr Ala Ala Arg Glu
          115          120          125

Leu Asp Pro Asp Glu Ala Lys His Leu Arg Ser Asp Leu Thr Arg Ala
          130          135          140

```

Arg Val Gly Tyr Glu Ile Glu Met Arg Thr Lys Val Lys Met Ile Met
 145 150 155 160

Gly Lys Leu Lys Ser Glu Gly Val Glu Ile Lys Val Thr Cys
 165 170

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 584 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

- (B) CLONE: 3B76

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CCTCCTCACTC CAGGCCAGCC AACAAAAGAA CCTACATTTA TTCCAGTGGT TGTTGGTCTT	60
TTGGACTCAA GTGGGAAAGA CATTACTCTT TCCTCTGTTC ATTATGATGG TACAGTGCAG	120
ACCATTTTCAG GCAGCAGCAC AATACTTCGA GTGACAAGAA ACAAGAAGAG TTTGTGTTTT	180
CTGATATACC AGAAAGACCT GTTCCGTCCC TATTTAGGGG ATTTCAGCCCC AGTTCGTGTT	240
GAAACTGATC TCTCTAATGA TGAATTATTC TTCTCTCTAG CACATGATTC AGATGAATTC	300
AATAGGTGGG AGGCCGGTCA AGTTCTGGCA AGAAAGCTGA TGCTGAACCTT AGTTTCTGAT	360
TTCCAGCAAA ATAAACCGTT GGCTCTAAAC CAAAATTTTG TGCAAGGTCT CGGCAGTGTG	420
CTTTCTGACT CAAGCTTGA CAAGGAATTT ATAGCCAAAG CAATAACACT ACCTGGGGAG	480
GGAGAGATAA TGGACATGAT GGCCGTGGCG GATCCTGATG CTGTTTCATGC TGTTAGAAAG	540
TTTGTACGAA AGCAGCTTGC ATCTGAACTT AAGGAGGAGC TTCT	584

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 283 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Arabidopsis thaliana*
- (vii) IMMEDIATE SOURCE:
 (B) CLONE: 3B76
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
- | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|
| Pro | Pro | Thr | Pro | Gly | Gln | Pro | Thr | Lys | Glu | Pro | Thr | Phe | Ile | Pro | Val | 1 | 5 | 10 | 15 |
| Val | Val | Gly | Leu | Leu | Asp | Ser | Ser | Gly | Lys | Asp | Ile | Thr | Leu | Ser | Ser | 20 | 25 | 30 | |
| Val | His | Tyr | Asp | Gly | Thr | Val | Gln | Thr | Ile | Thr | Gly | Ser | Ser | Thr | Ile | 35 | 40 | 45 | |
| Leu | Arg | Val | Thr | Lys | Lys | Gln | Glu | Glu | Phe | Val | Phe | Ser | Asp | Ile | Pro | 50 | 55 | 60 | |
| Glu | Arg | Pro | Val | Pro | Ser | Leu | Phe | Arg | Gly | Phe | Ser | Ala | Pro | Val | Arg | 65 | 70 | 75 | 80 |
| Val | Glu | Thr | Asp | Leu | Ser | Asn | Asp | Asp | Leu | Phe | Phe | Leu | Leu | Ala | His | 85 | 90 | 95 | |
| Asp | Ser | Asp | Glu | Phe | Asn | Arg | Trp | Glu | Ala | Gly | Gln | Val | Leu | Ala | Arg | 100 | 105 | 110 | |
| Lys | Leu | Met | Leu | Asn | Leu | Val | Ser | Asp | Phe | Gln | Gln | Asn | Lys | Pro | Leu | 115 | 120 | 125 | |
| Ala | Leu | Asn | Pro | Lys | Phe | Val | Gln | Gly | Leu | Gly | Ser | Val | Leu | Ser | Asp | 130 | 135 | 140 | |

Ser	Ser	Leu	Asp	Lys	Glu	Phe	Ile	Ala	Lys	Ala	Ile	Thr	Leu	Pro	Gly	145	150	155	160
Glu	Gly	Glu	Ile	Met	Asp	Met	Met	Ala	Val	Ala	Asp	Pro	Asp	Ala	Val	165	170	175	
His	Ala	Val	Arg	Lys	Phe	Val	Arg	Lys	Gln	Leu	Ala	Ser	Glu	Leu	Lys	180	185	190	
Glu	Glu	Leu	Lys	Ile	Val	Glu	Asn	Asn	Arg	Ser	Thr	Glu	Ala	Tyr	Val	195	200	205	
Phe	Asp	His	Ser	Asn	Met	Ala	Arg	Arg	Ala	Leu	Lys	Asn	Thr	Ala	Leu	210	215	220	
Ala	Tyr	Leu	Ala	Ser	Leu	Glu	Asp	Pro	Ala	Tyr	Met	Gly	Thr	Cys	Thr	225	230	235	240
Glu	Arg	Ile	Gln	Gly	Gly	His	Gln	Phe	Asp	Arg	Pro	Ile	Cys	Cys	Phe	245	250	255	
Gly	Thr	Leu	Ser	Gln	Asn	Pro	Gly	Lys	Thr	Arg	Glu	Arg	Thr	Phe	Leu	260	265	270	
Pro	Asp	Phe	Tyr	Glu	Gln	Val	Ala	Gly	Thr	Ile						275	280		

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 534 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: *Arabidopsis thaliana*
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: 4A5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

ACCAGGAGGG GAAAAAGTCT TACCCCATGG ACATCCCGGG GATTGAGTGT TACCCGAAAA 60
 GGATGAAGAA TGGTATTCTT CCGTCGTGGA CCCCATGCAC CCATTGGGAA AGCCGTGTGG 120
 CGTTTTCTTT CAGGGATGAT AGAAAAGTGC TCCCTTGGGA TGGAAAGGAG GAGCCTTTAC 180
 TGGTAGTGGC CGATAGGGTG AGGAATGTTG TGGAGGCTGA TGACGGGTAT TATCTCGTGG 240
 TGGCTGAGAA CGGACTTAAG CTAGAGAAAG GATCAGATTT GAAGGCGAGA GAGGTGAAGG 300
 AGAGTTTAGG GATGGTTGTT TTGGTGGTGA GGCCGCCAAG AGAAGATGAT GATGATTGGC 360
 AGACAAGTCA TCAGAACTGG GACTGAATTA ATAGAATCAA TACTCATATG CTGTAACTGA 420
 TTACGGAGTC ATCATGGTCA TGTAAAATTT TTGGATAAAG GTGGTAACTT TTTGTTCTAA 480
 GATACAATCA GAAACAGAGC AATATTTTTT TCTAAAAAAA AAAAAAAAAA AAAA 534

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*

(vii) IMMEDIATE SOURCE:

- (B) CLONE: 4A5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Met Asp Ile Pro Gly Ile Glu Cys Tyr Pro Lys Arg Met Lys Asn Gly
 1 5 10 15
 Ile Pro Pro Ser Trp Thr Pro Cys Thr His Trp Glu Ser Arg Val Ala
 20 25 30

Phe Ser Phe Arg Asp Asp Arg Lys Val Leu Pro Trp Asp Gly Lys Glu
 35 40 45
 Glu Pro Leu Leu Val Val Ala Asp Arg Val Arg Asn Val Val Glu Ala
 50 55 60
 Asp Asp Gly Tyr Tyr Leu Val Val Ala Glu Asn Gly Leu Lys Leu Glu
 65 70 75 80
 Lys Gly Ser Asp Leu Lys Ala Arg Glu Val Lys Glu Ser Leu Gly Met
 85 90 95
 Val Val Leu Val Val Arg Pro Pro Arg Glu Asp Asp Asp Asp Trp Gln
 100 105 110
 Thr Ser His Gln Asn Trp Asp
 115

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

- (B) CLONE: primer V6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ATGCTTTGCA TAACTTTGAG G

21

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: primer T7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

AATACGACTC ACTATAG

17